Reduction to Generalized Hessenberg Form and Inverse Spectral Problems*

Leonid Faybusovich

Division of Applied Sciences
Pierce Hall, Room G8
Harvard University
Cambridge, Massachusetts 02138

Submitted by Paul A. Fuhrmann

ABSTRACT

Lie-algebraic generalizations of Hessenberg matrices are considered. We address the questions under what conditions an element of the classical Lie algebra can be reduced by a similarity transformation to the generalized Hessenberg form. We completely solve an inverse spectral problem for symmetric generalized Hessenberg matrices related to classical Lie algebras.

1. INTRODUCTION

Hessenberg matrices play a prominent role in various problems of linear algebra. We mention only that the possibility of the reduction of a square matrix to the Hessenberg form by a similarity orthogonal transformation is very important for practical implementations of QR-like algorithms.

In the past few years there has been a considerable interest in the developing of QR-type algorithms which preserve various symmetry properties of a given initial matrix (see e.g. [5, 14] and references therein). In particular, R. Byers [4] (motivated by control problems) developed a QR-type algorithm for Hamiltonian matrices which is based on the Iwasawa decomposition in the symplectic group (see also [3] for other algorithms of this type).

*This work was supported in part by the National Science Foundation under grant CDR-85-00108.

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

R. Byers [4] also found an analogue of Hessenberg matrices for this situation. The same type of generalized Hessenberg matrices appear in a quite different context in [7]. Partially motivated by this example, several authors [1, 5, 6, 12] conjectured that the right analogue of Hessenberg matrices for an arbitrary semisimple real split Lie algebra $\mathscr{L}$ can be described as follows. Let

$$\mathscr{L} = \left(\sum_{\alpha \in \Delta} \mathbf{R} g_{\alpha} \right) \oplus \mathscr{A}$$

be a root-space decomposition of $\mathscr{L}$ ($\mathscr{A}$ is a fixed Cartan subalgebra of $\mathscr{L}$ such that all elements ada, $a \in \mathscr{A}$, have real spectrum). Denote by Δ_s the set of positive simple roots (relative to some Weyl chamber). Then elements of the subset

$$H_{\mathscr{L}} = \left(\sum_{\alpha \in \Delta_s} \mathbf{R}^* g_{-\alpha} \right) \oplus \left(\sum_{\alpha \in \Delta^+} \mathbf{R} g_{\alpha} \right) \oplus \mathscr{A}, \tag{1}$$

where Δ^+ is the set of positive roots, are called $\mathscr{L}$-Hessenberg. We arrive at the classical notion of the Hessenberg matrix for the case $\mathscr{L} = \mathrm{sl}_n$. If we take $\mathscr{L} = \mathrm{sp}_n$, $\mathscr{A} = \{\mathrm{diag}(\Lambda, -\Lambda),\ \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n),\ \lambda_i \in \mathbf{R}\}$, we obtain the Hamiltonian Hessenberg matrices introduced in [4]. In the present paper we address the following question. Let $G(\mathscr{L})$ be a connected Lie group corresponding to the Lie algebra $\mathscr{L}$. Given $\xi \in \mathscr{L}$, under what conditions does there exist $g \in G(\mathscr{L})$ such that $\mathrm{Ad}(g)\xi \in H_{\mathscr{L}}$? Let U be the maximal compact subgroup of $G(\mathscr{L})$. If the question posed above has a positive answer, one can find $g \in U$ such that $\mathrm{Ad}(g)\xi \in H_{\mathscr{L}}$ (this is a direct consequence of the Iwasawa decomposition). We also completely solve an inverse spectral problem for symmetric Hessenberg matrices related to classical Lie algebras.

We deliberately avoid any use of the structural theory of semisimple Lie algebras in the main text of this paper. Some comments connecting our constructions with abstract properties of semisimple Lie algebras can be found in the Appendix. Rather surprisingly, we found out that elementary tools from the realization theory of rational functions with internal symmetries are very useful for our purposes. But let us underscore once again that neither understanding this theory nor any preliminary knowledge of the structural properties of semisimple Lie algebras is necessary for reading this paper.

2. HESSENBERG MATRICES RELATED TO THE LIE ALGEBRA C_n

Throughout this paper we denote by (,) the standard scalar product in $\mathbf{R}^{2n}$. Let

$$J=\begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Here I_n is the $n\times n$ identity matrix. Consider a bilinear skew-symmetric nondegenerate form ω on $\mathbf{R}^{2n}$ (the standard symplectic structure):

$$\omega(x,y)=(x,Jy),\qquad x,y\in\mathbf{R}^{2n}.$$

Let us introduce the symplectic group Sp_n as the set of $2n\times 2n$ invertible matrices which preserve ω. Elements of the corresponding Lie algebra sp_n are usually called Hamiltonian matrices (see e.g. [9]). Thus $H\in\mathrm{sp}_n$ if and only if $\omega(Hx,y)+\omega(x,Hy)=0$, $x,y\in\mathbf{R}^{2n}$.

PROPOSITION 1. *Let*

$$H=\begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where A,B,C,D are real $n\times n$ matrices. Then $H\in\mathrm{sp}_n$ iff $D=-A^T$, $B=B^T$, $C=C^T$. Here A^T stands for the transpose of A.

The proof is straightforward (see e.g. [9]).

A vector subspace $V\subset\mathbf{R}^{2n}$ is called isotropic (relative to ω) if $\omega(x,y)=0$ for any $x,y\in V$. Let $V_1\subset V_2\subset\cdots\subset V_n$ be a sequence of isotropic vector subspaces in $\mathbf{R}^{2n}$ such that $\dim V_i=i$, $i=1,2,\dots,n$. We say that such a sequence forms a Lagrangian flag in $\mathbf{R}^{2n}$. We reserve the notation E for the flag $E_1\subset E_2\subset\cdots\subset E_n$, where $E_i=\mathrm{span}(e_1,\dots,e_i)$. Here $e_1,\dots e_i$ is the standard basis in $\mathbf{R}^n$.

DEFINITION 1. A matrix $H\in\mathrm{sp}_n$ is called a Hamiltonian Hessenberg matrix if the following conditions hold:

(i) the vectors $e_1,He_1,\dots,H^{2n-1}e_1$ are linearly independent over $\mathbf{R}$ (i.e., e_1 is a cyclic vector for H);

(ii) the flag $F(H)=(\mathbf{R}\,e_1\subset\mathrm{span}(e_1,\,He_1)\subset\cdots\subset\mathrm{span}(e_1,$ $He_1,\dots,H^{n-1}e_1))$ coincides with the standard flag E.

REMARK. Condition (i) of Definition 1 means that, in fact, we deal in this paper with so-called irreducible (or unreduced [2]) generalized Hessenberg matrices.

Denote the set of Hamiltonian Hessenberg matrices by H_C.

PROPOSITION 2. *Let* $H \in \mathrm{sp}_n$,

$$H = \begin{bmatrix} A & L \\ Q & -A^T \end{bmatrix}. \tag{2}$$

Then $H \in H_C$ *iff* A *is upper Hessenberg and* $Q = aE_{nn}$ *for some* $a \neq 0$.

REMARK. We will use the standard notation E_{ij} for the $n \times n$ matrix with the (i, j) entry equal to 1 and all other entries equal to 0.

Proof. If A is upper Hessenberg and $Q = aE_{nn}$, $a \neq 0$, then (i), (ii) of Definition 1 clearly hold. Conversely, from (ii) it follows that A is upper Hessenberg and if $Q = \|q_{ij}\|$, then $q_{ij} = 0$ for $j = 1, 2, \dots, n-1$. Since $Q = Q^T$, we obtain $q_{in} = 0$, $i = 1, 2, \dots, n-1$. Now (i) implies $q_{nn} \neq 0$. ■

Let G be the subgroup of Sp_n consisting of matrices having e_1 as an eigenvector. We will denote by char(H) the characteristic polynomial of a matrix H; $d(H)$ stands for $\omega(e_1, H^{2n-1}e_1)$.

THEOREM 1. *Let* $H_i \in H_C$, $i = 1, 2$. *Then* $H_2 = SH_1S^{-1}$ *for some* $S \in G$ *if and only if the following conditions hold*:

(i) $\mathrm{char}\, H_1 = \mathrm{char}\, H_2$;
(ii) $\mathrm{sign}\, dH_1 = \mathrm{sign}\, d(H_2)$.

Given any even polynomial $p(x) = x^{2n} + a_1x^{2n-2} + \cdots + a_n$, $a_i \in \mathbf{R}$, *and a nonzero real* a, *there always exists* $H \in H_C$ *such that* $d(H) = a$, $\mathrm{char}(H) = p$.

REMARK. Let G_C denote the normalizer of H_C under the adjoint action of G on sp_n. Theorem 1 can be reformulated as follows: $H_C / G_C \approx \mathbf{R}^n \times \{\pm 1\}$. As is easily seen, $G_C = \{g \in \mathrm{Sp}_n : gE = E\}$.

Proof. Let (i),(ii) hold. Then $e_1, H_1e_1, \dots, H_1^{2n-1}e_1$ is a basis in $\mathbf{R}^{2n}$. Consider a linear operator $S: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ which is defined as follows: $SH_1^ie_1 = H_2^ie_1$, $i = 0, 1, \dots, 2n-1$. S is invertible, since $H_2 \in H_C$. Further,

$$\omega\left(SH_1^ie_1, SH_1^je_1\right) = \omega\left(H_2^ie_1, H_2^je_1\right) = (-1)^i\omega\left(e_1, H_2^{i+j}e_1\right). \tag{3}$$

By Definition 1(ii) we have $\omega(e_1, H_1^ke_1) = \omega(e_1, H_2^ke_1) = 0$ if $k \leqslant 2n-2$. Let $\operatorname{char} H_i = p$, $i = 1, 2$. Of course, $p(x) = p(-x)$ (see e.g. [9]). If $p(x) = x^{2n} + a_1x^{2n-2} + \cdots + a_n$, then we have

$$H_i^{2n+1} = -a_1H_i^{2n-1} - a_2H_i^{2n-3} - \cdots - a_nH_i,$$

i.e. $\omega(e_1, H_i^{2n+1}e_1) = -a_1\omega(e_1, H_i^{2n-1}e_1)$ [because $\omega(e_1, H_i^{2n-(2k+1)}e_1) = 0$ for $k = 1, 2, \dots,$ as was mentioned above]. By induction we prove that

$$\omega\left(e_1, H_i^{2n+2k+1}e_1\right) = p_k(a_1, \dots, a_n)d(H_i), \tag{4}$$

$k = 1, 2, \dots,$ where p_k is a polynomial function in variables $a_1, \dots, a_n$ which does not depend on i.

Now without loss of generality we can suppose $d(H_1) = d(H_2)$. Indeed, let

$$S(a) = \begin{bmatrix} aI_n & 0 \\ 0 & a^{-1}I_n \end{bmatrix}, \qquad a \in \mathbf{R}^*.$$

We have $d(S(a)^{-1}HS(a)) = a^2d(H)$, $S(a) \in G$. Hence, by (ii) one can always find $a \in \mathbf{R}$ such that $d(S(a)^{-1}H_1S(a)) = d(H_2)$. If $d(H_1) = d(H_2)$, then (4) implies $\omega(e_1, H_1^ke_1) = \omega(e_1, H_2^ke_1)$ for all k. By (3) we have $S \in \mathrm{Sp}_n$ and consequently $S \in G$. Conversely, let $H_2 = SH_1S^{-1}$ for some $S \in G$. Then, of course, $\operatorname{char} H_2 = \operatorname{char} H_1$, $\operatorname{sign} d(H_1) = \operatorname{sign} d(H_2)$. Given any $p(x) = x^{2n} + a_1x^{2n-2} + \cdots + a_n$, $a_i \in \mathbf{R}$, $a \in \mathbf{R}^*$, we now construct $H \in H_C$ such that $\operatorname{char} H = p$, $d(H) = a$.

Consider the vector space P_{2n} of real polynomials of degree not greater than $2n-1$. Then P_{2n} has a basis $1, x, \dots, x^{2n-1}$. Consider a linear operator H in P_{2n}, such that $Hx^i = x^{i+1}$, $i = 1, 2, \dots, 2n-1$, $Hx^{2n-1} = -a_1x^{2n-2} - a_2x^{2n-4} - \cdots - a_n$. Introduce real numbers h_i, $i = 0, 1, \dots,$ as follows. Set $h_{2k} = 0$, $k = 0, 1, \dots,$ $h_i = 0$, $0 \leqslant i \leqslant 2n-2$, $h_{2n-1} = a$, and further define inductively

$$h_{2n+(2k+1)} = -a_1 h_{2n+(2k-1)} - a_2 h_{2n+(2k-3)} - \cdots - a_n h_{2k+1}, \qquad (5)$$

$k = 0, 1, \ldots$. We now introduce a bilinear form ω' on P_{2n} as follows: $\omega'(x^i, x^j) = (-1)^i h_{i+j}$.

Observe, first, that ω' is skew-symmetric. Indeed, $\omega'(x^j, x^i) = (-1)^j h_{i+j}$. If $i+j$ is even, then $h_{i+j} = 0$. If $i+j$ is odd, then $(-1)^j = -(-1)^i$. In both cases $\omega'(x^i, x^j) = -\omega'(x^j, x^i)$. We now verify that $\omega'(Hx^i, x^j) = -\omega'(x^i, Hx^j)$. If $i, j < 2n-1$, then $\omega'(Hx^i, x^j) = \omega'(x^{i+1}, x^j) = (-1)^{i+1} h_{i+j+1}$ and $\omega(x^i, Hx^j) = (-1)^i h_{i+j+1} = -\omega'(Hx^i, x^j)$. Let, say, $i = 2n-1$, $j < 2n-1$. We have $\omega'(Hx^{2n-1}, x^j) = -a_1 h_{2n-2+j} - a_2 h_{2n-4+j} - \cdots - a_n h_j$. On the other hand, $\omega'(x^{2n-1}, Hx^j) = -h_{2n+j}$. If j is even, then $\omega'(Hx^{2n-1}, x^j) = -\omega'(x^{2n-1}, x^{j+1}) = 0$. If j is odd, it is sufficient to refer to (5). Other cases are verified similarly.

We prove that ω' is nondegenerate. It is sufficient to verify that the matrix Ω with the entries $\omega'(x^i, x^j)$ is invertible. But

$$\Omega = \begin{bmatrix} 0 & T \\ -T & Z \end{bmatrix},$$

where T is a lower triangular (relative to the antidiagonal) matrix with elements $\omega(x^i, x^{2n-1-i}) = (-1)^i a$, $i = 0, 1, \ldots, n-1$, on the antidiagonal. Thus Ω is invertible. Since the n-dimensional vector subspace $\operatorname{span}(1, x, \ldots, x^{n-1})$ is isotropic, there exist vectors $f_0, \ldots, f_{n-1} \in P_{2n}$ such that $\omega'(x^i, f_j) - \delta_{ij} = \omega'(f_i, f_j) = 0$ (see e.g. [9]). In the (symplectic) basis $1, x, \ldots, x^{n-1}, f_0, \ldots, f_{n-1}$ the matrix of H is Hamiltonian. By the very definition of H, properties (i), (ii) of Definition 1 hold, $\operatorname{char} H = p$, and $d(H) = a$. ∎

REMARK. In the construction of a Hessenberg matrix with a given characteristic polynomial we used ideas from the realization theory of rational functions with internal symmetries. In fact, we defined a rational function

$$f(z) = \frac{h_1}{z^2} + \frac{h_3}{z^4} + \cdots$$

of McMillan degree $2n$ with the following additional property: $h_1 = h_3 = \cdots = h_{2n-3} = 0$. We further constructed a polynomial realization of f by a

Hamiltonian matrix H (see in this respect [8], where a much more general situation is considered) in the form

$$f(z) = \omega\left(e_1, (zI_{2n} - H)^{-1} e_1\right).$$

The matrix H is Hessenberg due to the imposed conditions. This method is pretty general and works for all classical Lie algebras.

One can define Hessenberg Hamiltonian matrices over $\mathbf{C}$ as in Definition 1 (see also [9]). Let $H_C(\mathbf{C}), G(\mathbf{C})$ be corresponding objects over $\mathbf{C}$.

THEOREM 2. *Let H_i, $i = 1,2$, be in $H_C(\mathbf{C})$. Then $H_2 = SH_1S^{-1}$ for some $S \in G(\mathbf{C})$ iff* $\operatorname{char} H_1 = \operatorname{char} H_2$. *Given any polynomial $p(x) = x^{2n} + a_{1x}^{2n-2} + \cdots + a_n$ with $a_i \in \mathbf{C}$, there always exists $H \in H_C(\mathbf{C})$ such that* $\operatorname{char} H = p$.

A proof of Theorem 2 is similar to that of Theorem 1 but easier.

REMARK. Let $H \in H_C$ have the form (2), where $Q = aE_{nn}$. Then $\operatorname{sign} d(H) = (-1)^n \operatorname{sign} a$.

COROLLARY 1. *Let $H \in \mathrm{sp}_n(\mathbf{C})$. The following statements are equivalent:*

(i) *There exists $S \in \mathrm{Sp}_n(\mathbf{C})$ such that $SHS^{-1} \in H_C(\mathbf{C})$.*
(ii) *H has a cyclic vector.*

Proof. Of course, (i) implies (ii). Conversely, given $H \in \mathrm{sp}_n(\mathbf{C})$, by Theorem 2 there exists $H_1 \in H_C(\mathbf{C})$ such that $\operatorname{char} H = \operatorname{char} H_1$. This implies that H_1, H have the same Jordan normal form (since both have cyclic vectors). Hence, there exists $S \in \mathrm{Sp}_n(\mathbf{C})$ such that $SHS^{-1} = H_1$ (see [11]). ■

COROLLARY 2. *Let $H \in \mathrm{sp}_n$ have no purely imaginary eigenvalues. The following statements are equivalent:*

(i) *There exists $S \in \mathrm{Sp}_n$ such that $SHS^{-1} \in H_C$.*
(ii) *There exists $O \in \mathrm{Sp}_n \cap O(2n, \mathbf{R})$ such that $OHO^{-1} \in H_C$.*
(iii) *H has a cyclic vector.*

REMARK. $O(2n, \mathbf{R})$ stands for the group of orthogonal transformations in $\mathbf{R}^{2n}$.

Proof. The implication (i) $\Rightarrow$ (iii) is obvious. Given $H \in \mathrm{sp}_n$ with a cyclic vector, by Theorem 1 there exists $H_1 \in H_C$ such that $\operatorname{char} H = \operatorname{char} H_1$.

Hence, H_1, H have the same Jordan normal form. Since H has no purely imaginary eigenvalues, there exists $S \in \mathrm{Sp}_n$ such that $SHS^{-1} = H_1$ (see [11]). This proves the implication (iii) $\Rightarrow$ (i).

Clearly (ii) implies (i). Given $S \in \mathrm{Sp}_n$, one can find $O \in O(2n, \mathbf{R}) \cap \mathrm{Sp}_n$ and $R \in \mathrm{Sp}_n$ with the following properties: $RE = E$, $S = RO$. (This is the Iwasawa decomposition in Sp_n [9]. The elements $R \in \mathrm{Sp}_n$ such that $RE = E$ form a subgroup B called a Borel subgroup.) It is clear by Definition 1 that $RH_C R^{-1} = H_C$ for any $R \in B$. Thus if $SHS^{-1} \in H_C$, then $OHO^{-1} \in H_C$. ∎

Corollary 3. *Any symmetric real Hamiltonian matrix with a cyclic vector can be reduced by an orthogonal, symplectic similarity transformation to the Hamiltonian Hessenberg form.*

Proof. Observe that a symmetric Hamiltonian matrix with a cyclic vector does not have zero eigenvalues. Indeed, such an eigenvalue would have an even multiplicity [9]. The result follows from Corollary 2. ∎

We show now that in the real case the existence of a cyclic vector is not sufficient for a reduction to the Hessenberg form in sp_n. Suppose that $H \in \mathrm{sp}_n$ has a simple purely imaginary spectrum

$$\{\pm i\omega_1, \pm i\omega_2, \ldots, \pm i\omega_n\}, \qquad 0 < \omega_1 < \omega_2 < \cdots < \omega_n. \tag{6}$$

Then [9] $\mathbf{R}^{2n} = V_1 \oplus V_2 \oplus \cdots \oplus V_n$, where V_i are H-invariant two-dimensional subspaces, pairwise orthogonal relative to ω. We call V_i positive (negative) if the restriction of the quadratic form $\omega(x, Hx)$ to V_i is positive (negative) definite; we set $\varepsilon_i = +1$ if V_i is positive, $\varepsilon_i = -1$ otherwise. As is well known (see e.g. [9]), two Hamiltonian matrices with the same spectrum (6) are similar by a real symplectic transformation iff the signs ε_i of the corresponding vector subspaces coincide. Of course, each matrix with the spectrum (6) has a cyclic vector (the minimal and the characteristic polynomials coincide in this case).

Proposition 3. *Suppose that $H \in H_C$ has a spectrum* (6). *Then $\varepsilon_i = (-1)^{i-1} \operatorname{sign} d(H)$, $i = 1, 2, \ldots, n$.*

Proof. Let $e_1 = x_1 + x_2 + \cdots + x_n$, $x_i \in V_i$. Then by Definition 1

$$\omega(x_1, H^{2k+1}x_1) + \cdots + \omega(x_n, H^{2k+1}x_n) = 0,$$

$$k = 0,1,\ldots,n-2;\ \omega\left(x_1, H^{2n-1}x_1\right) + \cdots + \omega\left(x_n, H^{2n+1}x_n\right) = d(H).$$

But $H^{2k+1}x_i = (-\omega_i^2)^k H x_i$, since $H^2 + \omega_i^2 I_{2n} = 0$ on V_i. If we denote $\omega(x_i, Hx_i)$ by y_i, we arrive at the following linear system:

$$\begin{aligned} &\sum_{i=1}^{n} \left(-\omega_i^2\right)^k y_i = 0, \qquad k = 0,1,\ldots,n-2, \\ &\sum_{i=1}^{n} \left(-\omega_i^2\right)^{n-1} y_i = d(H). \end{aligned} \tag{7}$$

Let $p(x) = (x+\omega_1^2)(x+\omega_2^2)\cdots(x+\omega_n^2)$ and $p_i(x) = p(x)/(x+\omega_i^2)$. Then (7) has the following solution: $y_i = d(H)/p_i(-\omega_i^2)$, i.e., $\operatorname{sign} y_i = (-1)^{i-1}\operatorname{sign} d(H)$, $i = 1,2,\ldots,n$. ■

Remark. The technique used in the proof of Proposition 3 enables one to obtain an explicit criterion for reducibility to a Hamiltonian Hessenberg form. We will address this question elsewhere. On the other hand, we don't consider any practical algorithm for the actual reduction to generalized Hessenberg forms. See in this respect [2].

We now turn to the analysis of the structure of isospectral sets in H_C. We denote by H_C^s the set of symmetric matrices in H_C. The elements of H_C^s should be considered as analogues of tridiagonal symmetric matrices. In fact, under an appropriate choice of the orthonormal basis in $\mathbf{R}^{2n}$ they can be simultaneously reduced to this form. Introduce a map f (which goes back to Moser [13]),

$$f : \mathrm{gl}(2n, \mathbf{R}) \to \mathrm{Rat}(\mathbf{R}), \qquad F(H)(z) = \left(e_1, (zI_{2n} - H)^{-1} e_1\right). \tag{8}$$

Here Rat(**R**) stands for the manifold of scalar rational functions with real coefficients.

Proposition 4. *Let* $H_1, H_2 \in H_C^s$, $\operatorname{sign} d(H_1) = \operatorname{sign} d(H_2)$. *Then* $f(H_1) = f(H_2)$ *if and only if* $H_2 = SH_1S^{-1}$, $S = \operatorname{diag}(K,K)$, $K = \operatorname{diag}(\varepsilon_1,\ldots,\varepsilon_n)$, $\varepsilon_i = \pm 1$, $i = 1,2,\ldots,n$.

Proof. Let $f(H_1) = f(H_2)$. As in the proof of Theorem 1, define $S : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ as follows: $SH_1^i e_1 = H_2^i e_1$, $i = 0,1,\ldots,2n-1$. We have $(SH_1^i e_1, SH_1^j e_1) = (H_2^i e_1, H_2^j e_1) = (e_1, H_2^{i+j} e_1) = (H_1^i e_1, H_1^j e_1)$ [since $H_i^T = H_i$, $f(H_1) = f(H_2)$].

This yields $S \in O(2n, \mathbf{R})$. Since $H_i \in H_C$, we have $F(H_i) = E$ (see Definition 1). By construction $SF(H_1) = F(H_2)$, which implies $S = \operatorname{diag}(K_1, K_2)$, $K_1 = \operatorname{diag}(\delta_1, \dots, \delta_n)$, $K_2 = \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_n)$, $\delta_i, \varepsilon_i = \pm 1$. We also have $SH_1S^{-1} = H_2$. Let

$$H_1 = \begin{bmatrix} A & L \\ L & -A \end{bmatrix}, \qquad A = \|a_{ij}\|, \quad L = \|q_{ij}\|, \tag{9}$$

where A is symmetric and tridiagonal with $a_{i+1,i} \neq 0$, $i = 1, 2, \dots, n-1$; $Q = aE_{nn}$ (see Proposition 2). Then

$$SH_1S^{-1} = \begin{bmatrix} K_1AK_1 & K_1LK_2 \\ K_2LK_1 & -K_2AK_2 \end{bmatrix} = H_2.$$

Since $H_2 \in H_C$, we obtain $\varepsilon_i\varepsilon_{i+1} = \delta_i\delta_{i+1}$, $i = 1, 2, \dots, n-1$. Moreover, $\varepsilon_n\delta_n = 1$ because $\operatorname{sign} d(H_1) = \operatorname{sign} d(H_2)$. (Here we have used the Remark after Theorem 2.) This yields $\varepsilon_i = \delta_i$. ∎

Denote by H_C^{s+} the subset of H_C^s consisting of the matrices (9) with $a_{i+1,i} > 0$, $d(H) > 0$. By Proposition 4 we have

Corollary. *The map f restricted to H_C^{s+} is injective.*

Let

$$\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \qquad M = \operatorname{diag}(\Lambda, -\Lambda), \tag{10}$$

where $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. Denote by $H_C^{s+}(\Lambda)$ the set of matrices from H_C^{s+} with the fixed spectrum $\{\pm\lambda_1, \pm\lambda_2, \dots, \pm\lambda_n\}$.

Theorem 3. *The image $f(H_C^{s+}(\Lambda))$ is the set of rational functions of the form*

$$g(z) = \sum_{i=1}^{l} \left(\frac{l_i}{z - \lambda_i} + \frac{m_i}{z + \lambda_i} \right), \tag{11}$$

$$m_i = c\frac{\gamma_i^2}{l_i}, \quad \gamma_i = \frac{1}{p_i(\lambda_i)\lambda_i}, \qquad c = \frac{1 - \sum_{i=1}^n l_i}{\sum_{i=1}^n (\gamma_i^2 / l_i)}. \tag{12}$$

Here $p_i(x) = (x - \lambda_1^2)(x - \lambda_2^2) \cdots (x - \lambda_n^2)/(x - \lambda_i^2)$, *and* $l = (l_1, \dots, l_n)$ *runs over the set* $\{l \in \mathbf{R}^n : l_1 + l_2 + \cdots + l_n < 1,\ l_i > 0,\ i = 1, 2, \dots, n\}$.

REMARK. Thus $H_C^{s+}(\Lambda)$ is isomorphic with $\mathbf{R}^n$. This result follows from general Lie-algebraic considerations [10].

Proof. In this proof P stands for the set $H_C^{s+}(\Lambda)$. Let $H \in P$. Then [9] there exists a symplectic, orthogonal transformation S such that $SHS^{-1} = M$ [see (10)]. Denote $S^{-1}e_1$ by r. We have $\omega(r, M^{2k-1}r) = \omega(e_1, H^{2k-1}e_1) = 0$, $k = 1, 2, \dots, n-1$ (see Definition 1). Further, $\omega(r, M^{2n-1}r) = d(H)$. If $r = [x, y]^T$, $x, y \in \mathbf{R}^n$, we arrive at the following linear system:

$$\begin{aligned} &\lambda_1^{2k-1}x_1y_1 + \lambda_2^{2k-1}x_2y_2 + \cdots + \lambda_n^{2k-1}x_ny_n = 0, \qquad k = 1, 2, \dots, n-1, \\ &\lambda_1^{2n-1}x_1y_1 + \lambda_2^{2n-1}x_2y_2 + \cdots + \lambda_n^{2n-1}x_ny_n = \frac{-d(H)}{2}. \end{aligned} \tag{13}$$

Hence, $x_iy_i = -\gamma_i d(H)/2$, where γ_i is defined in (12). This solution of (13) is obtained by the inversion of the Vandermonde matrix. On the other hand, $f(H) = (r, (zI_{2n} - M)^{-1}r)$ is given by (11) with $l_i = x_i^2$, $m_i = y_i^2$. Since $S \in O(2n, \mathbf{R})$, we have $x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 = 1$. Using our solution to (13), we obtain the relations (12) with $c = d(H)^2/4$.

Conversely, suppose a rational function (11) with conditions (12) is given. Observe that the vectors $r, Mr, \dots, M^{2n-1}r$, where $r = [x, y]^T$, $x_i = l_i^{1/2}$, $y_i = -m_i^{1/2}$, $i = 1, 2, \dots, n$, are linearly independent. Indeed,

$$\begin{aligned} &r \wedge Mr \wedge \cdots \wedge M^{2n-1}r \\ &\quad = x_1x_2 \cdots x_ny_1 \cdots y_n\lambda_1 \cdots \lambda_n \prod_{1 \leqslant i < j \leqslant n} (\lambda_i^2 - \lambda_j^2) e_1 \wedge e_2 \wedge \cdots \wedge e_{2n} \neq 0. \end{aligned}$$

Further, the flag F, $F_i = \operatorname{span}(r, Mr, \dots, M^{i-1}r)$, $i = 1, 2, \dots, n$, is Lagrangian, since (12) implies that x_iy_i, $i = 1, 2, \dots, n$, are solutions to (13) with $d(H) = 2c^{1/2}$. Let $h_1, \dots, h_n$ be an orthonormal basis in F_n such that $\operatorname{span}(h_1, \dots, h_i) = F_i$, $i = 1, 2, \dots, n$, $h_1 = r$. Consider a matrix

$$T = [h_1, h_2, \dots, h_n, -Jh_1, -Jh_2, \dots, -Jh_n],$$

where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

As is easily seen, $T \in \mathrm{Sp}_n \cap O(2n, \mathbf{R})$. Let $H = T^{-1}MT$. By the very definition of T we have $TE_i = F_i$, $i = 1, 2, \ldots, n$, $Te_1 = r$. Thus $H^k e_1 = T^{-1}M^k r \in E_i$, $k = 0, 1, \ldots, i-1$. In other words, $F(H) = E$ (see Definition 1). Besides, $e_1 \wedge He_1 \wedge \cdots \wedge H^{2n-1}e_1 = (\det T)^{-1}(r \wedge Mr \wedge \cdots \wedge M^{2n-1}r) \neq 0$. Thus $H \in H_C^s$. By the construction $f(H) = g$, $d(H) = \omega(r, M^{2n-1}r) = 2c^{1/2} > 0$. One can always find $S = \mathrm{diag}(K, K)$, $K = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_n)$, $e_i = \pm 1$, such that $SHS^{-1} \in P$. Since $d(SHS^{-1}) = d(H)$, $f(SHS^{-1}) = f(H)$, we are done. ∎

REMARK. Given "spectral data" (l, m, Λ) subject to the conditions (10), (11), one can recover (a unique) element $H \in H_C^{s+}(\Lambda)$ as in the proof of Theorem 3.

3. HESSENBERG MATRICES RELATED TO THE LIE ALGEBRA D_n

Throughout this section $n \geqslant 2$.

Let

$$J' = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}.$$

Consider a bilinear form B on $\mathbf{R}^{2n}$:

$$B(x, y) = (x, J'y), \qquad x, y \in \mathbf{R}^{2n}.$$

As in the case of $\mathbf{C}_n$, we introduce the Lie group O_n of invertible matrices which preserve B. The corresponding Lie algebra (notation: o_n) consists of the matrices H such that

$$B(Hx, y) + B(x, Hy) = 0, \qquad x, y \in \mathbf{R}^{2n}.$$

PROPOSITION 5. *Let*

$$H = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where A, B, C, D *are real* $n \times n$ *matrices. Then* $H \in o_n$ *iff* $D = -A^T$, $C = -C^T$, $B = -B^T$.

The proof is a straightforward calculation.

The notions of an isotropic subspace and a Lagrangian flag are defined as in the symplectic case.

DEFINITION 2. A matrix $H \in o_n$ is called D-Hessenberg if the following conditions hold:

(i) the vector e_1 is cyclic for H;

(ii) the vector spaces $(e_1, He_1, \dots, H^{i-1}e_1)$ coincide with the vector spaces E_i, $i = 1, 2, \dots, n-1$,

REMARK. Observe the difference in conditions (ii) of Definitions 1, 2.

PROPOSITION 6. *Let* $H \in o_n$,

$$H = \begin{bmatrix} A & L \\ Q & -A^T \end{bmatrix}. \tag{14}$$

Then H *is D-Hessenberg if and only if* A *is upper Hessenberg and* $Q = a(E_{n,n-1} - E_{n-1,n})$ *for some* $a \neq 0$.

Proof. If H has the form (14), then clearly H is C-Hessenberg. Conversely, if H satisfies conditions (i), (ii) of Definition 2, then A is upper Hessenberg and the first $n-2$ columns of Q are equal to zero. Let $He_{n-1} = a_1 e_1 + \cdots + a_n e_n + b_1 f_1 + \cdots + b_n f_n$, where we use the notation $f_i = e_{n+i}$, $i = 1, 2, \dots, n$. Then $B(e_i, He_{n-1}) = -B(He_i, e_{n-1}) = 0$ for $i = 1, 2, \dots, n-2$, since $He_i \in \operatorname{span}(e_1, He_1, \dots, H^i e_1) = E_{i+1}$. This implies $b_i = 0$ for $i = 1, 2, \dots, n-2$. Further, $B(e_{n-1}, He_{n-1}) = -B(e_{n-1}, He_{n-1}) = 0$, i.e. $b_{n-1} = 0$. Finally, $b_n \neq 0$ (otherwise, $e_1 \wedge He_1 \wedge \cdots \wedge H^{n-1}e_1 = 0$). The required form of Q follows from $Q^T = -Q$. ■

We introduce some notation. Let H_D denote the set of D-Hessenberg matrices, G be the subgroup of O_n consisting of the matrices which have e_1 as an eigenvector, and $d(H)$ stand for $B(e_1, H^{2(n-1)}e_1)$.

THEOREM 4. *Let $H_1, H_2 \in H_D$. Then $H_2 = SH_1S^{-1}$ for some $S \in G$ if and only if the following conditions hold*:

(i) $\operatorname{char} H_1 = \operatorname{char} H_2$;
(ii) $\operatorname{sign} d(H_1) = \operatorname{sign} d(H_2)$.

Proof. The proof is quite similar to that of Theorem 1 and will be only outlined here. Without loss of generality one can take $d(H_1) = d(H_2)$. Consider a linear operator $S: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ which is defined as follows: $SH_1^i e_1 = H_2^i e_1$, $i = 0, 1, \dots, 2n-1$. Clearly, $S \in \mathrm{GL}(2n, \mathbf{R})$. We should check $B(SH_1^i e_1, SH_1^j e_1) = B(H_1^i e_1, H_1^j e_1)$. This can be done exactly as in the proof of Theorem 1. Indeed, let $\operatorname{char} H_i = x^{2n} + a_1 x^{2n-2} + \cdots + a_n$, $i = 1, 2$. Then $H_i^{2n} = -a_1 H_i^{2(n-1)} - \cdots - a_n I_{2n}$. Observe that $B(e_1, H_i^k e_1) = 0$, $0 \leqslant k \leqslant 2(n-2)$. We can argue now as in the proof of Theorem 1. ∎

Let

$$p(x) = x^{2n} + a_1 x^{2(n-1)} + \cdots + a_n, \qquad a_i \in \mathbf{R}, \quad \in \mathbf{R}^*. \tag{15}$$

Introduce the following sequence h_i: $h_{2k-1} = 0$, $k = 1, 2, \dots$; $h_i = 0$, $0 \leqslant i \leqslant 2(n-2)$, $h_{2(n-1)} = a$, $h_{2n+2k} = -a_1 h_{2n-2+2k} - a_2 h_{2n-4+2k} - \cdots - a_n h_{2k}$, $k = 0, 1, \dots$.

THEOREM 5. *There exists $H \in H_D$ such that* $\operatorname{char} H = p$, $d(H) = a$ *if and only if*

$$\operatorname{sign}\det\begin{bmatrix} h_2 & h_4 & \cdots & h_{2n} \\ h_4 & h_6 & \cdots & h_{2n+2} \\ \vdots & \vdots & & \vdots \\ h_{2n} & h_{2n+2} & \cdots & h_{4n-2} \end{bmatrix} = (\operatorname{sign} a)^n (-1)^{n(n-1)/2}. \tag{16}$$

Proof. Let $H \in H_D$ be such that $\operatorname{char} H = p$, $d(H) = a$. We verify (16). Consider the vector subspaces $V = \operatorname{span}(e_1, H^2 e_1, \dots, H^{2n-2} e_1)$, $W = HV$. Since $H \in H_D$, we clearly have $V \oplus W = \mathbf{R}^{2n}$. Moreover, $W = V^0$ (where by V^0 we denote the orthogonal compliment of V in $\mathbf{R}^{2n}$ relative to B). In particular, V, W are nondegenerate relative to B. Consider the vector subspace $L = \operatorname{span}(e_1, He_1, \dots, H^{n-2} e_1)$. By Definition 2, L is isotropic relative to B.

At this point we need some elementary properties of bilinear forms [11]. Let X be a nondegenerate subspace. We call a vector subspace Y of X

positive (negative) if $B(x,x)>0$ (<0), $x\in Y\setminus\{0\}$. The dimension p (m) of a maximal positive (negative) vector subspace in X depends only on X. We have $p+m=\dim X$, $\operatorname{sign} X = p-m$. Denote by $\operatorname{ind} X$ the dimension of a maximal isotropic subspace in X. As is well known, $\operatorname{ind}(X)=\min(m,p)$. We now consider the following situations: $n=2k$, $n=2k+1$.

(a) $n=2k$. Observe that $\dim(L\cap V)=k$, $\dim(L\cap W)=k-1$. This implies $\operatorname{sign} V=0$. Since $0=\operatorname{sign}\mathbf{R}^{2n}=\operatorname{sign} V+\operatorname{sign} W$, we have $\operatorname{sign} W=0$, i.e. $\operatorname{ind} W=k$. Let us write down the Gram matrix Γ of $B|W$ in the basis $He_1, H^3e_1,\dots,H^{2n-1}e_1$. We have $\Gamma=\|\gamma_{ij}\|$, $\gamma_{ij}=B(H^{2i-1}e_1,H^{2j-1}e_1)$, $i,j=1,2,\dots,n$. Now $\gamma_{ij}=-B(e_1,H^{2(i+j-1)}e_1)=-h_{2(i+j-1)}$. Since $\operatorname{ind} W=0$, we must have $\operatorname{sign}\det\Gamma=(-1)^k$. But this is equivalent to (16), since $n=2k$ is even.

(b) $n=2k+1$. In this case $\dim(V\cap L)=k$, $\dim(W\cap L)=k$. We have the following possibilities for $\operatorname{sign} V$: $\operatorname{sign} V=1$, $\operatorname{sign} V=-1$. Let $\operatorname{sign} V=1$; then of course $\operatorname{sign} W=-1$, which implies $\det\Gamma=(-1)^{k+1}$. Consider the Gram matrix Γ' of $B|V$ in the basis $e_1,H^2e_1,\dots,H^{2(n-1)}e_1$. We have $\Gamma=\|\gamma'_{ij}\|$, $\gamma'_{ij}=B(H^{2i}e_1,H^{2j}e_1)=h_{2(i+j)}$, $i,j=0,1,\dots,n-1$. But $h_i=0$, $0\leqslant i\leqslant 2(n-2)$. This implies

$$\Gamma'=\begin{bmatrix} 0 & & & h_{2(n-1)} \\ & & h_{2(n-1)} & * \\ & \vdots & * & * \\ h_{2(n-1)} & \cdots & * & * \end{bmatrix},$$

i.e. $\det\Gamma'=(-1)^{n(n+1)/2}h^n_{2(n-1)}$, i.e. $\operatorname{sign}\det\Gamma'=(-1)^{n(n+1)/2}(\operatorname{sign} a)^n$. On the other hand, since $\operatorname{sign} V=1$, we must have $\det\Gamma'=(-1)^k$. Comparing this with $\operatorname{sign}\det\Gamma=(-1)^{k+1}$, we arrive at (16). The case $\operatorname{sign} V=-1$ is analyzed similarly. Inversely, if (16) holds one can construct $H\in H_D$ with a given $\operatorname{char} H$, $d(H)$ as in the proof of Theorem 1. We leave the details to the reader. ■

REMARK. One can define D-Hessenberg matrices over $\mathbf{C}$. Let $H_D(\mathbf{C}), G(\mathbf{C})$ be corresponding objects over $\mathbf{C}$.

THEOREM 6. *Let $H_1,H_2\in H_D(\mathbf{C})$. Then $H_2=SH_1S^{-1}$ for some $S\in G(\mathbf{C})$ iff* $\operatorname{char} H_1=\operatorname{char} H_2$. *Given any polynomial* $p(x)=x^{2n}+a_1x^{2n-2}$

$+ \cdots + a_n$, $a_i \in \mathbf{C}$, *construct the matrix*

$$\mathscr{H} = \begin{bmatrix} h_2 & h_4 & \cdots & h_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ h_{2n} & h_{2n+2} & \cdots & h_{4n-2} \end{bmatrix} \tag{17}$$

as in Theorem 5. *There exists* $H \in H_D(\mathbf{C})$ *such that* $\operatorname{char} H = p$ *iff* $\det \mathscr{H} \neq 0$.

COROLLARY 1. *Suppose that* $H \in H_D(\mathbf{C})$, $\det \mathscr{H} \neq 0$, *where* $\mathscr{H}$ *is constructed from* $\operatorname{char} H$ *as in Theorem* 5. *The following statements are equivalent*:

(i) *there exists* $S \in O_n(\mathbf{C})$ *such that* $SHS^{-1} \in H_D(\mathbf{C})$;
(ii) *H has a cyclic vector.*

COROLLARY 2. *Suppose that* $H \in o_n$ *and has no purely imaginary eigenvalues. If the condition* (16) *holds, then the following statements are equivalent*:

(i) *there exists* $S \in O_{2n}$ *such that* $SHS^{-1} \in H_D$;
(ii) *H has a cyclic vector.*

We now turn to the description of isospectral manifolds of symmetric D-Hessenberg matrices (notation: H_D^s). At this point we need to choose an orientation in $\mathbf{R}^{2n}$. We say that a basis $r_1, \dots, r_{2n}$ in $\mathbf{R}^{2n}$ is positively oriented if $r_1 \wedge \cdots \wedge r_{2n} = ce_1 \wedge e_2 \wedge \cdots \wedge e_{2n}$ for some $c > 0$. Given a Lagrangian flag $F = F_1 \subset F_2 \subset \cdots \subset F_n$, an orthonormal set of vectors $r_1, \dots, r_n$ such that $F_i = \operatorname{span}(r_1, \dots, r_i)$, $i = 1, 2, \dots, n$, is defined uniquely up to possible changes in signs (the only possible choices of such orthonormal sets are $\pm r_1, \dots, \pm r_n$). We say that F is positively oriented if $r_1, \dots, r_n$, $J'r_1, \dots, J'r_n$ is a positively oriented basis in R^{2n}. Clearly, this definition is correct (i.e. does not depend on a choice of signs of r_i).

PROPOSITION 7. *Let* $F_1 \subseteq F_2 \subset \cdots \subset F_{n-1}$, $\dim F_i = i$, $i = 1, 2, \dots, n-1$, *be a "partial" isotropic flag* (*i.e.*, F_{n-1} *is an isotropic subspace*). *Then there exist exactly two Lagrangian flags* $N = (N_1 \subset \cdots \subset N_n)$ *such that* $N_i = F_i$, $i = 1, 2, \dots n-1$. *Exactly one of these Lagrangian flags is positively oriented.*

Proof. Let $r_1, \dots, r_{n-1}$ be an orthonormal in F_{n-1} such that $F_i = \operatorname{span}(r_1, \dots, r_{n-1})$, $i = 1, 2, \dots, n-1$. One can find $r_n \in \mathbf{R}^{2n}$ such that $(r_n, r_i) = 0$, $B(r_n, r_i) = 0$, $i = 1, 2, \dots, n$, $(r_n, r_n) = 1$ (see e.g. [6]). Set $f_i = J'r_i$,

$i=1,2,\dots,n$. Then $B(r_i,f_j)-\delta_{ij}=B(f_i,f_j)=0$. Suppose that N, $\dim N=n$, is any isotropic subspace containing F_{n-1}, and $r\in N\setminus F_{n-1}$. Let $r=a_1r_1+\cdots+a_nr_n+b_1f_1+\cdots+b_nf_n$, $a_i,b_i\in\mathbf{R}$. Since $B(r,r_i)=0$, $i=1,2,\dots,n-1$, we have $b_i=0$, $i=1,2,\dots,n-1$. Further, $B(r,r)=a_nb_n=0$. The case $a_n=b_n=0$ would imply $r\in F_{n-1}$, which contradicts the choice of r. If $a_n\neq 0$, $b_n=0$, then $N=\operatorname{span}(r_1,\dots,r_n)$. Otherwise, $N=\operatorname{span}(r_1,\dots,r_{n-1},f_n)$. Now

$$r_1\wedge r_2\wedge\cdots\wedge r_n\wedge f_1\wedge f_2\wedge\cdots\wedge f_n$$
$$=-r_1\wedge r_2\wedge\cdots\wedge r_{n-1}\wedge f_n\wedge f_1\wedge\cdots\wedge f_{n-1}\wedge r_n. \qquad \blacksquare$$

Denote by $H_D^{s+}(\Lambda)$ the set of matrices H from H_D^s with a fixed spectrum $\{\lambda_1,\dots,\lambda_n,-\lambda_1,\dots,-\lambda_n\}$ with the following properties:

(i) $d(H)>0$;

(ii) there exists $O\in O_n\cap O(2n,\mathbf{R})$, $\det O=1$, such that $H=OMO^{-1}$ [see (10)];

(iii) If H has the form (14) with $A=\|a_{ij}\|$, then $a_{i+1,i}>0$, $i=1,2,\dots,n-1$.

THEOREM 7. *The map f, described in* (8), *establishes a one-one correspondence between $H_D^{s+}(\Lambda)$ and the set of rational functions g of the form*

$$g(z)=\sum_{i=1}^{n}\left(\frac{l_i}{z-\lambda_i}+\frac{m_i}{z+\lambda_i}\right), \tag{18}$$

$$m_i=\frac{\gamma_i^2}{l_i}c,\ \gamma_i=\frac{1}{p_i(\lambda_i)},\quad i=1,2,\dots,n,\qquad c=\frac{1-\sum_{i=1}^n l_i}{\sum_{i=1}^n\left(\gamma_i^2/l_i\right)}. \tag{19}$$

Here $p_i(x)=(x-\lambda_1^2)(x-\lambda_2^2)\cdots(x-\lambda_n^2)/(x-\lambda_i^2)$ and $l=(l_1,\dots,l_n)$ runs over the set $\{l\in\mathbf{R}^n: l_1>0,\dots,l_n>0,\ l_1+l_2+\cdots+l_n<1\}$.

Proof. In this proof we will use the following notation: P for $H_d^{s+}(\Lambda)$, and U for $\{O\in O_n\cap O(2n,\mathbf{R}):\det O=1\}$. Let $H\in P$, $O\in U$ be such that $H=OMO^{-1}$ [see (10)]. Denote $O^{-1}e_1$ by r. We have by Definition 2: $B(r,M^{2k-2}r)=B(e_1,H^{2k-2}e_1)=0$, $k=1,2,\dots,n-1$. Further, $B(r,M^{2n-2}r)=d(H)$. If $r=[x,y]^T$, $x,y\in\mathbf{R}^n$, we arrive at the following linear system:

$$\lambda_1^{2k}x_1y_1 + \cdots + \lambda_n^{2k}x_ny_n = 0, \qquad k = 0,1,\ldots,n-2,$$
$$\lambda_1^{2n-2}x_1y_1 + \cdots + \lambda_n^{2n-2}x_ny_n = \frac{d(H)}{2}. \tag{20}$$

Hence, $x_iy_i = \gamma_i d(H)/2$, $i = 1,2,\ldots,n$. On the other hand, $f(H) = (r,(zI_{2n} - H)^{-1}r)$ is given by (18) with $l_i = x_i^2$, $m_i = y_i^2$. Since $x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 = 1$, we obtain (19) with $c = d(H)^2/4$. Conversely, if for a rational function (18) the conditions (19) hold, then the vector space $F_{n-1} = \operatorname{span}(r, Mr,\ldots,M^{n-2}r)$, $r = [x,y]^T$, $x_i = l_i^{1/2}$, $y_i = m_i^{1/2}$, $i = 1,2,\ldots,n$, is isotropic. Observe that $r \wedge Mr \wedge \cdots \wedge M^{2n-1}r \neq 0$, whence $\dim F_{n-1} = n-1$. By Proposition 7 there exists a unique n-dimensional isotropic subspace F_n containing F_{n-1} such that the flag $F_1 \subset F_2 \subset \cdots \subset F_n$, where $F_i = \operatorname{span}(r, Mr,\ldots,M^{i-1}r)$, $i = 1,2,\ldots,n-1$, is positively oriented. Let $r_1,\ldots,r_n$ be an orthonormal basis in F_n such that $\operatorname{span}(r_1,\ldots,r_i) = F_i$, $i = 1,2,\ldots,n$, $r_1 = r$. Consider the matrix $T = [r_1,\ldots,r_n, J'r_1,\ldots,J'r_n]$. Clearly $T \in U$ ($\det T = 1$, since the flag F is positively oriented by the construction). As in the proof of Theorem 3, we see that $H = T^{-1}MT$ belongs to H_D^s, $f(H) = g$, $d(H) > 0$. This implies (see the proof of Theorem 3) that $g = f(H')$ for some $H' \in P$.

It remains to prove that f is injective on P. Let $H = OMO^{-1} \in P$, $O \in U$. We prove that O is "almost" determined by $f(H)$. Indeed, if $r = O^{-1}e_1$, then r is determined by $f(H)$ up to the multiplication by a matrix $D = \operatorname{diag}(\varepsilon_1,\ldots,\varepsilon_{2n})$, $\varepsilon_i = \pm 1$, from the left [see (19)]. Moreover, from (20) we have $\varepsilon_i = \varepsilon_{n+i}$, $i = 1,2,\ldots,n$ [since $d(H) > 0$]. Hence, the positively oriented flag $F = (F_1 \subset \cdots \subset F_n)$, $\phi_i = \operatorname{span}(r, Mr,\ldots,M^{i-1}r)$, $i = 1,2,\ldots,n-1$, which coincides with $O^{-1}E$ is determined uniquely up to the action by D from the left. This defines O uniquely up to multiplication by D_1, D_2 from the left and from the right, where D_i are of the form $\operatorname{diag}(K,K)$, $K = \operatorname{diag}(\varepsilon_1,\ldots,\varepsilon_n)$, $\varepsilon_i = \pm 1$. In other words, if $f(H_1) = f(H_2)$, $H_i \in P$, then $H_2 = D_1H_1D_1^{-1}$. But if

$$H_i = \begin{bmatrix} A_i & L_i \\ L_i & -A_i \end{bmatrix}, \qquad i = 1,2,$$

$A_i = \|a_{st}^{(i)}\|$, then we must have $a_{j+1,j}^{(i)} > 0$, $j = 1,2,\ldots,n-1$, $i = 1,2$, i.e. $D_1 = I_{2n}$ or $D_1 = -I_{2n}$. In both cases $H_2 = H_1$. ∎

4. CONCLUDING REMARKS

In the present paper we have considered Lie-algebraic generalizations of Hessenberg matrices for classical Lie algebras C_n, D_n. It was shown that elements of these Lie algebras by the adjoint action of the corresponding Lie groups can be reduced to the Hessenberg form provided some mild additional conditions are imposed. Unlike the classical A_n case, the existence of a cyclic vector is not, in general, sufficient for the existence of such a reduction. Nevertheless, for almost all situations arising in control theory this reduction is possible. We also completely solved analogues of inverse spectral problems for symmetric Hessenberg matrices related to classical Lie algebras.

We briefly describe a natural abstract Lie-algebraic setting for our constructions in the Appendix.

APPENDIX

In this Appendix we use some elementary facts about semisimple Lie algebras (see e.g. [10]). Let $\mathscr{L}$ be a semisimple real split Lie algebra. We denote by $G(\mathscr{L})$ the corresponding connected Lie group. Fix a Cartan decomposition

$$\mathscr{L} = K \oplus P. \tag{A.1}$$

Let, further, $A \subset P$ be a Cartan subalgebra and $C \subset A$ be a Weyl chamber. In this way we obtain a root-space decomposition of $\mathscr{L}$:

$$\mathscr{L} = \left(\sum_{\alpha \in \Delta} \mathbf{R} e_\alpha \right) \oplus A.$$

Denote by θ the Cartan involution corresponding to the Cartan decomposition (A.1). Take $e_{-\alpha} = -\theta e_{-\alpha}$, $\alpha \in \Delta^+$ (the set of positive roots relative to C). Set $g_\alpha = e_\alpha + e_{-\alpha}$, $\alpha \in \Delta^+$; Δ_s stands for the set of simple positive roots.

DEFINITION 3. Elements of the set

$$H_{\mathscr{L}} = \left(\sum_{\alpha \in \Delta^+} \mathbf{R} e_\alpha \right) \oplus A \oplus \left(\sum_{\alpha \in \Delta_s} \mathbf{R}^* e_{-\alpha} \right)$$

are called Hessenberg elements of $\mathscr{L}$.

DEFINITION 4. Elements of the set

$$H_{\mathscr{L}}^s = H_{\mathscr{L}} \cap P$$

are called symmetric Hessenberg.

It is clear that

$$H_{\mathscr{L}}^s = \left(\sum_{\alpha \in \Delta_s} \mathbf{R}^* g_\alpha \right) \oplus A.$$

In the present paper we address the following questions:

(i) Given $\xi \in \mathscr{L}$, find $g \in G(\mathscr{L})$ such that $\mathrm{Ad}(g)\xi \in H_{\mathscr{L}}$.

(ii) One can show [10] that each element of $H_{\mathscr{L}}^s$ is regular, i.e., if we denote by U the connected Lie subgroup of $G(\mathscr{L})$ corresponding to the Lie algebra K, then for each $\xi \in H_{\mathscr{L}}^s$ there exists $O \in U$ such that $\varphi(\xi) = \mathrm{Ad}(O)\xi \in C$ [the element $\varphi(\xi) \in C$ is defined uniquely]. Given $c \in C$, denote by $H_{\mathscr{L}}^s(c)$ the set $\varphi^{-1}(c)$. The second main question which we address is this: Describe the sets $H_{\mathscr{L}}^s(c)$ in terms of c and some additional parameters.

Let $B = \Sigma\{\mathbf{R} e_\alpha : \alpha \in \Delta^+\} \oplus A$. Denote by $O(B)$ the connected subgroup in $G(\mathscr{L})$ with the Lie algebra B. Each element $g \in G(\mathscr{L})$ admits a unique decomposition $g = QR$ with $Q \in U$, $R \in G(B)$ (the Iwasawa decomposition). Let $\xi \in \mathscr{L}$ and $C(\xi)$ be the centralizer of ξ in $G(\mathscr{L})$, i.e. $C(\xi) = \{g \in G(\mathscr{L}) : \mathrm{Ad}(g)\xi = \xi\}$. Fix some element $\psi(\xi) \in C(\xi)$. Define by induction a sequence $\xi_1, \xi_2, \ldots$ in $\mathscr{L}$ as follows. Set $\xi_1 = \xi$. Take the Iwasawa decomposition $\psi(\xi_1) = Q_1 R_1$. Set $\xi_2 = \mathrm{Ad}(Q_1^{-1})\xi_1$. Take the Iwasawa decomposition of $\psi(\xi_2) = \mathrm{Ad}(Q_1^{-1})\psi(\xi_1) = Q_2 R_2$, etc. The obtained sequence yields the iterates of the QR algorithm (with a constant multiple shift) corresponding to the Iwasawa decomposition [5].

PROPOSITION 8. *if $\xi_1 \in H_{\mathscr{L}}$, then all iterates ξ_i, $i = 2, 3, \ldots$ belong to $H_{\mathscr{L}}$.*

Proof. Let we have already proved that $\xi_1, \ldots, \xi_k \in H_{\mathscr{L}}$ and $\psi(\xi_k) \in C(\xi_k)$. Then $\xi_{k+1} = \mathrm{Ad}(Q_k^{-1})\xi_k$, $\psi(\xi_k) = Q_k R_k$, where $Q_k \in U$, $R_k \in G(B)$. We clearly have $\psi(\xi_{k+1}) \in C(\xi_{k+1})$ and $\xi_{k+1} = \mathrm{Ad}(R_k)\,\xi_k$. But $\mathrm{Ad}(G(B))\, H_{\mathscr{L}} \subset H_{\mathscr{L}}$. ∎

Problem (i) has a very transparent meaning in this context: given $\xi \in \mathscr{L}$, reduce it first to the Hessenberg form and then apply the QR algorithm. For example, R. Byers [4] considers the QR algorithm based on the Iwasawa decomposition in Sp_n In [4] ψ is taken to be a usual Cayley transform. It is proposed to reduce a Hamiltonian matrix to the C-Hessenberg form as an initial step of the QR algorithm.

Problem (ii) appears in various contexts (e.g., its solution can be used for a construction of action-angle variables for Toda flows related to classical Lie algebras [10]).

In this paper we have considered simple Lie algebras C_n, D_n. In both cases a Cartan subalgebra A is taken to be the set of diagonal matrices $M = \mathrm{diag}(\Lambda, -\Lambda)$, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. In the case of C_n the Weyl chamber C is chosen to consist of the matrices M such that $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. In the case of D_n we have $C = \{M \in A : \lambda_1 > \lambda_2 > \cdots > |\lambda_n| > 0\}$. The remaining classical case B_n poses no additional problems and can be analyzed by methods developed in this paper.

I would like to express my sincere gratitude to the referee for very helpful comments.

REFERENCES

1 G. S. Ammar and C. F. Martin, The geometry of matrix eigenvalue methods, *Acta Appl. Math.* 5:239–278 (1986).

2 G. S. Ammar and V. Mehrmann, On Hamiltonian and symplectic Hessenberg forms, *Linear Algebra Appl.*, to appear.

3 A. Bunse Gerstner, Matrix factorizations for symplectic QR-like methods, *Linear Algebra Appl.* 83:49–77 (1986).

4 R. Byers, A Hamiltonian QR-algorithm. *SIAM J. Sci. Statist. Comput.* 7:212–229 (1986).

5 L. Faybusovich, QR-type factorizations, the Yang-Baxter equation, and an eigenvalue problem of control theory, *Linear Algebra Appl.* 122–124:943–971 (1989).

6 L. Faybusovich, QR-algorithm and generalized Toda flows, *Ukranian Math. J.* 41(7):944–952 (1989).

7 P. A. Fuhrmann, On the Hamiltonian structure in the computation of singular values for a class of Hankel operators, preprint.

8 P. A. Fuhrmann, On symmetric rational transfer functions, *Linear Algebra Appl.* 50:167–250 (1983).

9 A. Givental, Symplectic geometry, in *Dynamical Systems IV* (V. I. Arnold and S. P. Novikov, Eds.), Springer-Verlag, New York, 1990.

10 B. Kostant, The solution to a generalized Toda lattice and representation theory, *Adv. in Math.* 34:195–338 (1979).

11 A. I. Maltsev, *Foundations of Linear Algebra*, Freeman, San Francisco, 1963.
12 F. De Mari and M. A. Shayman, Lie algebraic generalizations of Hessenberg matrices and the topology of Hessenberg varieties, Preprint.
13 J. Moser, Finitely many mass points on the line under the influence of an exponential potential—an integrable system, in *Lecture Notes in Phys.* 38 (J. Moser, Ed.), Springer-Verlag, Berlin, 1975.
14 D. Watkins and L. Elsner, Self-similar flows, *Linear Algebra Appl.* 110:213–242 (1988).

Received 16 October 1990; final manuscript accepted 17 January 1991